Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 30 (2012) 913 – 921

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference on Communication Technology and System Design 2011

Video Mining using LIM Based Clustering and Self Organizing Maps

C. Lakshmi Devasena[a], M.Hemalatha[b], a*

[a,b]*Department of Computer Science, Karpagam University, Coimbatore-641021,India*
Karpagam University,Coimbatore-641021,India

Abstract

Video mining has grown as an energetic research area and given incremental concentration in recent years due to impressive and rapid raise in the volume of digital video databases. The aim of this research work is to find out new objects in videos. This work proposes a novel approach for video mining using LIM based clustering technique and self organizing maps to recognize novelty in the frames of video sequence. The proposed work is designed and implemented on MATLAB. It is tested with the sample videos and provides promising results. And it is suitable for day to day video mining applications and object detection systems including remote video surveillance in defense for national and international border tracking.

Key words: Discrete Cosine Transform; Histogram; Lorenz Information Measure; Object Detection; Self Organizing Maps; Video Mining

1. Introduction

The field of image and video retrieval has been grown as a dynamic research area and has been waged more and more attention in recent years [1]. Due to availability of video in digital form, many organizations use video as content and store them in databases for their later analysis and utilization. Consequently, researchers are increasingly applying computer vision techniques to analyze the video before indexing. The concept of Object detection is used to find novelties in video. Object detection is the recognition of novel or strange data or indication that a machine learning system is not known during training. It is one of the primary requirements of a superior classification or identification system since occasionally the test data includes information about objects that were not identified at the time of training

* C. Lakshmi Devasena, Tel.: +91-9245861298
E-mail address: devaradhe2007@gmail.com

1877-7058 © 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.01.945

the model. Object detection can be based on the difference between actual and perceived external stimulus. Markus Timuska et al. [2] state: “For novelty detection, a description of normality is learnt by choosing an appropriate model to a set of normal examples, and previously unobserved patterns are then tested by comparing their novelty score (as defined by the model) against some threshold”. The main challenge in object detection is the definition of an appropriate model of known data and a threshold with which outliers can be detected for a given application. A number of methodologies for outlier detection have been developed over the years and they have their own drawbacks [3]. This work addresses some of the issues. Theoretically, a neural network can be trained to recognize the data directly. The goal of this research work is to derive a new method for novelty detection in video sequences using SOM. Detection of strange video events is significant for both consumer video applications such as commercial message detection and sports highlights extraction and also for video surveillance applications.

2. Literature Review

Xingquan Zhu et al. [4] suggested the Video association mining to construct knowledge based video indexing structure. A database management framework for video and strategies for video content structure and events mining is proposed by Xingquan Zhua et al. [5] using video mining techniques for efficient video database indexing, management and access. Nagia M. Ghanem et al. [6] proposed the petri net model for video surveillance system. Multiscale statistical models are used by Lexing Xie et al. [7] to produce algorithms for unsupervised mining of structures in video. Ajay Divakaran et al. [8] describes how video mining can be used for pattern recovery by comparing Unsupervised and Supervised approach for highlights extraction. New data mining framework for the extraction of goal events in soccer videos using combined multimodal analysis and decision tree logic is proposed by Shu-Ching Chen et al. [9]. Probablistic graphical models are used by Rong Yan et al. [10] to mine the relationship between video concepts. Major problems in automated video surveillance are discussed by Anthony R. Dick and Michael J. Brooks [11]. Several common object detection methods are proposed by Yilmaz, A et al. [12]. Several methods for performing background subtraction have been proposed by Massimo Piccardi [13]. Novelty Detection using parametric approaches are discussed by Chow[14], Hansenet et al. [15], Foggia et al. [16], Funera et al. [17] and Wei et al. [18] using various methods like Optimal threshold for rejection, Role of classifier confidence, MES architecture with Bayesian Combining Rule, Multiple Thresholds and Injection of anomalies. Novelty Detection using Non-parametric approaches are discussed by Hellman[19], Odin and Addison [20], Yang and Liu [21, 22], Duda et al. [23] and Dasgupta and Majundar [24] using Nearest Neighbour Classifier, k-nearest neighbor algorithm, Linear Least Squares Fit (LLSF), Parzen Windows method and Multi-dimensional data. Novelty Detection using clustering approaches are discussed by Bezdek et al. [25] and Pizzi et al. [26].

3. Methodology and Design

The design of the proposed research involves some of the image processing and statistical data mining and machine learning techniques.

3.1. Techniques used in Proposed Method

3.1.1. Histogram

Gray level / Color histogram shows the frequency of occurrence of each gray level / color in the image versus the gray level itself and provides a global description of the appearance of the image. The histogram with gray levels in the range [0, L-1] of a digital image is a discrete function.

$$P(rk) = nk / n \tag{1}$$

where,

rk - gray level *K*

nk - number of pixels in the image with the gray level *rk*.

n - total number of pixels contained in the image.

K = 0, 1, 2… *L*-1.

L = 256.

P(*rk*) gives an estimate of the probability of occurrence of gray level *rk*.

3.1.2. Lorenz Information Measure

Lorenz Information Measure (LIM) widely used in economics. Rorvig [27] was the first to suggest use of general features extracted from the images for retrieval and represented as LIMs. The Lorenz Information Measure (LIM) (P1,….,Pn) is defined to be the area under the Lorenz information curve. The area of LIM Ca is greater than the area of LIM Cb. Clearly, 0 < = LIM (P1,……., Pn) < = 0.5. If the probability vector is (P1,…….,Pn), then LIM (P1,…….,Pn) can be measured by the first ordering Pi's, and then calculating the area under the piecewise linear curve. Because LIM (P1,…….,Pn) (which can be expressed as the sum of f(Pi), and f(Pi)) is a continuous convex function, LIM (P1,…….,Pn) is considered as an information measure. Spontaneously, the LIM can be considered as a universal content-based information measure. To calculate the area of histograms, the histogram intervals are arranged from low to high, and the resulting off-diagonal shape measured through differentiation.

The formula to compute LIM is

1. *P0 = 0*
2. *W*= Width (Interval of histogram) has to be equal distribution
3. *0 <= LIM (P1, P2, …, Pn) <= 0.5.*
4. *LIM = {W*P1/2 + (W*P1 + (P2 – P1)/2) + … + (W*Pn-1 + W*(Pn – Pn-1)/2)}/2 * #ofpixels*

$$LIM = \{W * \sum^{n-1} Pi + W * (\sum^{n} Pi - Pi - 1) / 2)\} / 2 * \#ofpixels$$

$$LIM = \{W * \sum^{n-1} Pi + W * (\sum^{n} Pi - Pi - 1) / 2)\} / 2 * \#ofpixels \tag{2}$$

3.1.3. Discrete Cosine Transform

The Discrete Cosine Transform (DCT) can be used to Create feature based Image Profile. The DCT is a real domain transform which symbolizes the entire image as the coefficients of distinct frequencies of cosines (which are the source vectors for this transform). The DCT of the image is calculated by taking 8x8 blocks of the image, which are then transformed individually. The two dimensional DC Transform of an image gives the result matrix such that top left corner signifies lowest frequency coefficient whereas the bottom right corner is the highest frequency. The 1-D discrete cosine transform (DCT) is defined as

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x).\cos[(2x+1)u\pi / 2N] \quad (3)$$

In the same way, the inverse of the DCT is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cdot\cos\left[\frac{(2x+1)u\pi}{2N}\right] \quad (4)$$

where

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{for } u = 0 \\ \sqrt{2/N} & \text{for } u = 1,2,\ldots,N-1 \end{cases} \quad (5)$$

The equivalent 2-D DCT, and its inverse are defined as

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cdot\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cdot\cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (6)$$

and

$$f(x,y) = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \alpha(u)\alpha(v)C(u,v)\cdot\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cdot\cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (7)$$

The benefit of DCT is that it can be expressed without complex numbers. 2-D DCT is also separable (like 2-D Fourier transform), i.e. it can be obtained by two subsequent one dimensional DCT in the same way than Fourier transform [28].

3.1.4. Self Organizing Maps(SOM)

The SOM architecture is extensively different to multi- layer perceptrons and it is predominantly unsupervised techniques that recognize clusters in a data set, and moves the position of the neurons in feature space to represent these identified clusters. When a SOM is trained on common data, it will generate a kernel-based representation of normality which can be used for object detection. When training is complete and new data is passed, the Euclidean distance of output nodes representing data clusters can be thresholded to determine objects or strange entities present in the new data.

3.2. The LIM based Clustering and Self Organizing Maps Algorithm

The Steps involved in this algorithm are:

1. Open the Video to be trained and separate the content into RGB components (*iR*, *iG* and *iB*) and store it in separate Matrices.
2. Separate the video content into its HSV components (*iH*, *iS*, and *iV*) and store it in separate Matrices.
3. Find the LIM factor of the *iR*, *iG*, *iB*, *iH*, *iS*, and *iV*.
4. Find the DCT of the Matrices and store them separately in the corresponding Matrices *diR*, *diG*, *diB*, *diH*, *diS*, and *diV*.

5. Find the LIM factor *LdiR*, *LdiG*, *LdiB*, *LdiH*, *LdiS*, and *LdiV* of *diR*, *diG*, *diB*, *diH*, *diS*, and *diV*.
6. Cluster the LIM features of the sequence using SOM.
7. Open an Input video in which the objects to be detected and convert it to its RGB components and store it in separate Matrices.
8. Separate the input video in to its HSV components and store them in Separate Matrices.
9. Find the LIM factor of the *jR*, *jG*, *jB*, *jH*, *jS*, and *jV*.
10. Find the DCT of the Matrices and store them separately in the corresponding Matrices.
11. Find the LIM factor *LdjR*, *LdjG*, *LdjB*, *LdjH*, *LdjS*, and *LdjV* of *djR*, *djG*, *djB*, *djH*, *djS*, and *djV*.
12. Find the distance of the two LIM sets input video I and the trained video J from the Database.
13. The Distance

$d = \text{sum}\ ([LiR\ LiG\ LiB\ LiH\ LiS\ LiV\ LdiR\ LdiG\ LdiB\ LdiH\ LdiS\ LdiV] - [LjR\ LjG\ LjB\ LjH\ LjS\ LjV\ LdjR\ LdjG\ LdjB\ LdjH\ LdjS\ LdjV]\ )^2$

$d = \text{sum}\ [dr\ dg\ db\ dh\ ds\ dv\ ddr\ ddg\ ddb\ ddh\ dds\ ddv]^2$

$d = dr^2+dg^2+db^2+dh^2+ds^2+dv^2+ddr^2+ddg^2+ddb^2+ddh^2+dds^2+ddv^2$

14. Store the distance d in distance matrix *D*.
15. Repeat all the steps from 7 and find the distances of each and every frame image of input video in the image database.
16. Now the column matrix *D* will contain the distances of the image database from the input video.
17. Find the index of the lowest distance in the matrix *D*.
18. Display the corresponding video frame from the database which has strange object other than the normal content.

So in this proposed algorithm, to minimize the time, there will be two major steps.

- Create a neural network with x input neurons and k outputs.
- The same neural network is used to map the N number of patterns into k number of segments.

3.3. Model of the Proposed Work

The model of the proposed work is shown in the Fig. 1.

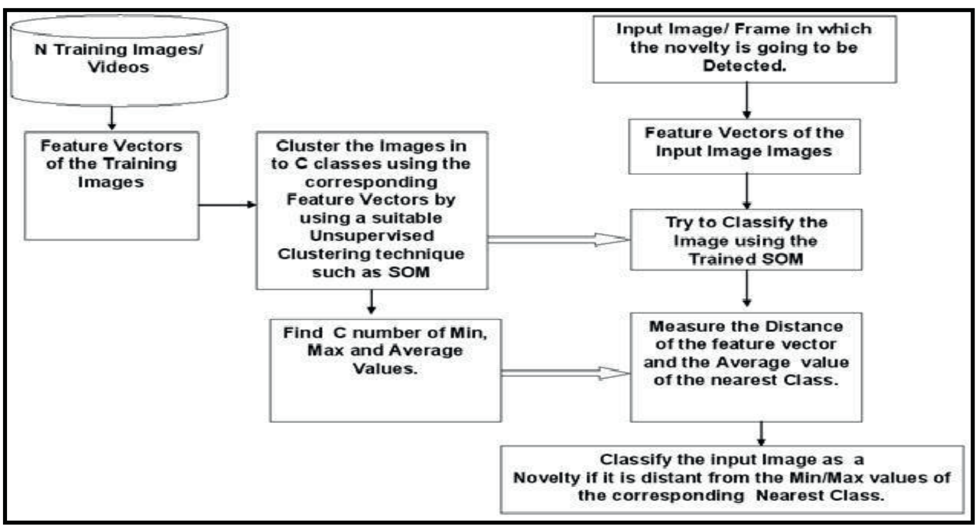

Fig. 1: Model of the proposed system

4. Results and Discussion

The Main Interface of the proposed work is shown in Fig. 2(a). First, the training video is selected (Fig. 2(b) and 3(a)). In the similar way the input video is selected. When we press ‘Train with video’, the training video is converted into indexed images i.e., the video is converted into array of frames. The features of the frames are extracted using LIM measure and clustered.

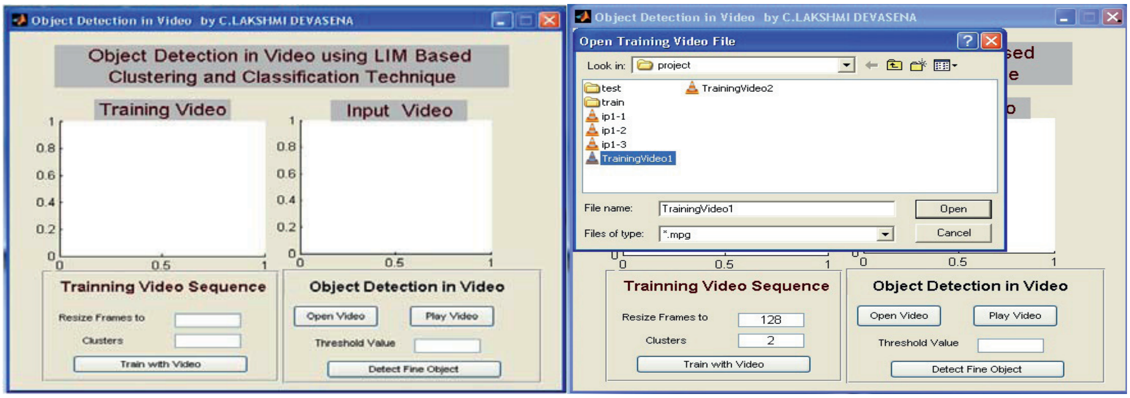

Fig. 2. (a) Main Interface; (b) Selecting Training Video

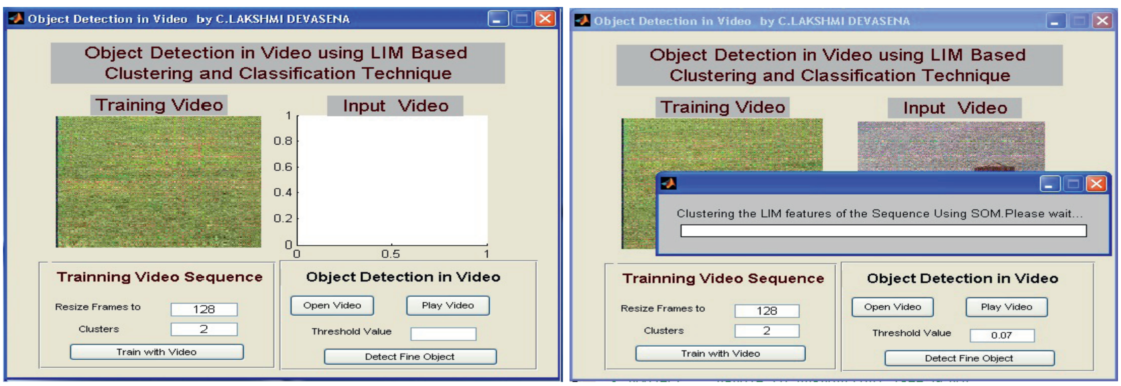

Fig. 3. (a) Input Video Selected; (b) Clusters LIM features of Training Video

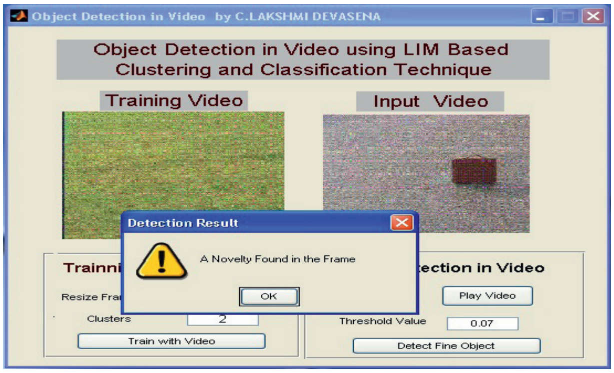

Fig. 4. Input Video Selected

By using Self organizing maps and the neural network, the training is performed (Fig. 3(b)). When we press 'Detect Fine Object' it convert the input video into indexed images and it extract the features of the image frames and compare it with the clusters of the training image and finds the novelty (Fig. 4). The features are selected by using the Lorenz information measure for the DCT features detected for the training and Input videos. For the sample video selected here the novelty has been found in an intermediate video frame at a threshold of 0.07. The threshold is the difference of the sum of squares of the LIM Features of the input video frame and the cluster centres of the SOM clustered using LIM features of the training video. Overall Results of the Novelty detection with three input Videos and two Training Videos are shown in Table1.

Table 1. Overall results of the Novelty Detection

S. No	Name of the Training Video File	Name of the Input Video File	Novelty Detected at Threshold
1	TrainingVideo1.mpg	ip1_1.mpg	0.17
2	TrainingVideo1.mpg	ip1_2.mpg	0.22
3	TrainingVideo2.mpg	ip1_1.mpg	0.22
4	TrainingVideo2.mpg	ip1_2.mpg	0.07
5	TrainingVideo2.mpg	ip1_3.mpg	0.05

5. Conclusion

The proposed LIM based clustering and self organizing maps technique for object detection framework has been successfully implemented and evaluated on Matlab. The performance of the LIM based clustering and SOM technique has been studied with different videos. Appropriate measures were formulated to evaluate the performance of the system. Results driven from the proposed system were noteworthy and comparable. While comparing with the number of false retrievals with the correct retrievals the proposed system seems to be achieved a performance level which will be is suitable for day to day video mining applications and object detection systems including remote video surveillance in defense for national and international border tracking. The video mining system derived from the LIM based clustering and SOM technique provided promising results. This kind of simple systems can be used in monitoring forest for fire or animal passing or human trespassing, where, normally, there will not be much change in background scene.

Acknowledgements

We thank Karpagam University for providing motivation and support for this work become success.

References

[1] C. Lakshmi Devasena, T.Sumathi, Dr. M. Hemalatha. Image Video, "Retrieval Technique: Grand Challenges and Trends" *Proceedings of National Conference on Applications of Data Mining in National Security (NCOADMINS 2010);* ISBN No: 987-93-80697-51-2, 2010, p. 96-105.

[2] Markus Timuska, Mike Lipsettb, Chris K. Mechefskec."Fault detection using transient machine signals" *ELSEVIER Mechanical Systems and Signal Processing;* 2008, p. 1724–1749.

[3] Leon Reznik, Gregory Von Pless, Tayeb Al Karim. "Application Testing of Novel Neural Network Structures" *Proceedings of the Workshop on Building Computational Intelligence and Machine Learning Virtual Organizations*; 2008.

[4] Xingquan Zhu, Xindong Wu, Ahmed K. Elmagarmid, Zhe Feng, Lide Wu. "Video Data Mining: Semantic Indexing and Event Detection from the Association Perspective" *IEEE Transactions on Knowledge And Data Engineering*; Volume 17, No 5, 2005.

[5] Xingquan Zhu, Walid G. Aref, Jianping Fan, Ann C. Catlin, Ahmed K. Elmagarmid. "Medical Video Mining for Efficient Database Indexing, Management and Access" *Proceedings of the 19th International Conference on Data Engineering (ICDE'03)*; 2003.

[6] Nagia M. Ghanem, David Doermann, Larry Davis, Daniel DeMenthon. "Mining Tool for Surveillance Video"

[7] Lexing Xie, Shih-Fu Changy, Ajay Divakaran, Huifang Sun. "Unsupervised Mining of Statistical Temporal Structures in Video" 2003.

[8] Ajay Divakaran, Kadir A. Peker, Shih-Fu Chang, Regunathan Radhakrishnan, Lexing Xie. "Video Mining: Pattern Discovery Versus Pattern Recognition" *Mitsubishi ElectricResearchLaboratories Inc.*; 2004.

[9] Shu-Ching Chen Mei-Ling Shyu Min Chen, Chengcui Zhang. A Decision Tree-based Multimodal Data Mining Framework for Soccer Goal Detection

[10] Rong Yan, Ming-yu Chen, Alexander Hauptmann. "Mining relationship between video concepts using probabilistic Graphical models"

[11] Anthony R. Dick, Michael J. Brooks. "Issues in Automated Visual Surveillance. 2004"

[12] Yilmaz, A., Javed, O., Shah, M. Object Tracking: "A Survey. *ACM Comput. Surv.*" Volume 38, No 4, Article 13, Dec. 2006.

[13] Massimo Piccardi. "Background subtraction techniques: A review" *IEEE International Conference on Systems, man and cybernets*; 2004.

[14] C.K. Chow."On optimum recognition error and reject tradeoff" *IEEE Transactions on Information Theory;* Volume IT-16, No. 1, January 1970, p. 41-46.

[15] L. K. Hansen, C. Liisberg, P. Salamon. "The error-reject tradeoff. *Open Systems and Information Dynamics*" Volume 4, 1997, p. 159-184.

[16] P. Foggia, C. Sansone, F. Tortorella, M. Vento. "Multiclassification: reject criteria for the Bayesian combiner" *Pattern Recognition* Volume 32, No. 8, 1999, p. 1435-1447.

[17] G. Fumera, F. Roli, G. Giacinto. "Reject option with multiple thresholds" *Pattern Recognition*; Volume 33, 2000, p. 2099-2101.

[18] F. Wei, M. Miller, S.J. Stolfo, L. Wenke, P.K. Chan. "Using artificial anomalies to detect unknown and known network intrusions" *Proc. IEEE International Conference on Data Mining (ICDM 2001)*; 2001, p. 123–130.

[19] M.E. Hellman. "The nearest neighbour classification with a reject option" *IEEE Transactions on Systems Science and Cybernetics*; Volume 6, No. 3, July 1970, p. 179-185.

[20] T. Odin, Addison D. "Novelty detection using neural network technology" *Proc. COMADEN conference*; 2000.

[21] Y. Yang, X. Liu. "A re-examination of text categorization methods". *Proc. ACM SIGIR Conference on Research and Development in Information Retrieval*; 1999, p. 42-49.

[22] Y. Yang, J. Zhang, J. Carbonell, C. Jin. "Topic-conditioned novelty detection" *International Conference on Knowledge Discovery and Data Mining*; July 2002.

[23] R.O. Duda, P.E. Hart, D.G. Stork. *Pattern classification*; Wiley, 2001.

[24] D. Dasgupta, N. S. Majumdar."Anomaly detection in multidimensional data using negative selection algorithm" *Proc. IEEE Conference on Evolutionary Computation, Hawai;*, May 2002, p. 1039-1044.

[25] J.C. Bezdek, R. Ehrlich, W. Full."FCM: the fuzzy c-means clustering algorithm" *Computers and Geosciences*; Volume 10, 1984, p. 191-203.

[26] N. J. Pizzi, R. A. Vivanco, R. L. Somorjai. "EvIdent: a functional magnetic resonance image analysis system" *Artificial Intelligence in Medicine*; Volume 21, 2001, p. 263-269.

[27] Mark Rorvig. "Image Retrieval by Content Measure Metadata Coding"

[28] S. Sakthivel, M. Rajaram. "Enhancing Clarity of Face Recognition to Augment the Accuracy Rate through Weighted Attribute Sets" *European Journal of Scientific Research*; ISSN 1450-216X, Volume 48, No.4, 2011, p.537-54.